Introduction

In this study, I explore a subset of MRI data from a longitudinal study focusing on patients with and without dementia. The dataset includes MRI scores and dementia status, and my goal is to analyze how these scores evolve over time using mixed-effects ANOVA and conduct a statistical power analysis for t-tests.

Research Questions

1. How do clinical dementia rating scores change among patients with and without dementia?
2. Is there a significant interaction between education level and dementia status on clinical dementia rating scores?
3. Does age influence the relationship between clinical dementia rating scores, gender, and dementia status?

Data Exploration

Before delving into the analysis, I explored the data to understand the distribution of clinical dementia rating scores by gender, dementia status, and education level. I created a new column ***education_group*** and grouped education levels by 0-10, 11-15, 16-20 for more meaningful analysis results. The EDA revealed trends in CDR scores across different variables between patients with and without dementia.

Descriptive Statistics by Demographic Factors
Calculated summary statistics (mean, median, standard deviation) of CDR scores grouped by gender, dementia status, and education level.

Visualization of CDR Distributions
Created boxplots (figure 1) and histograms (figure 2) to visualize the distribution of CDR scores by gender, dementia status, and education level.

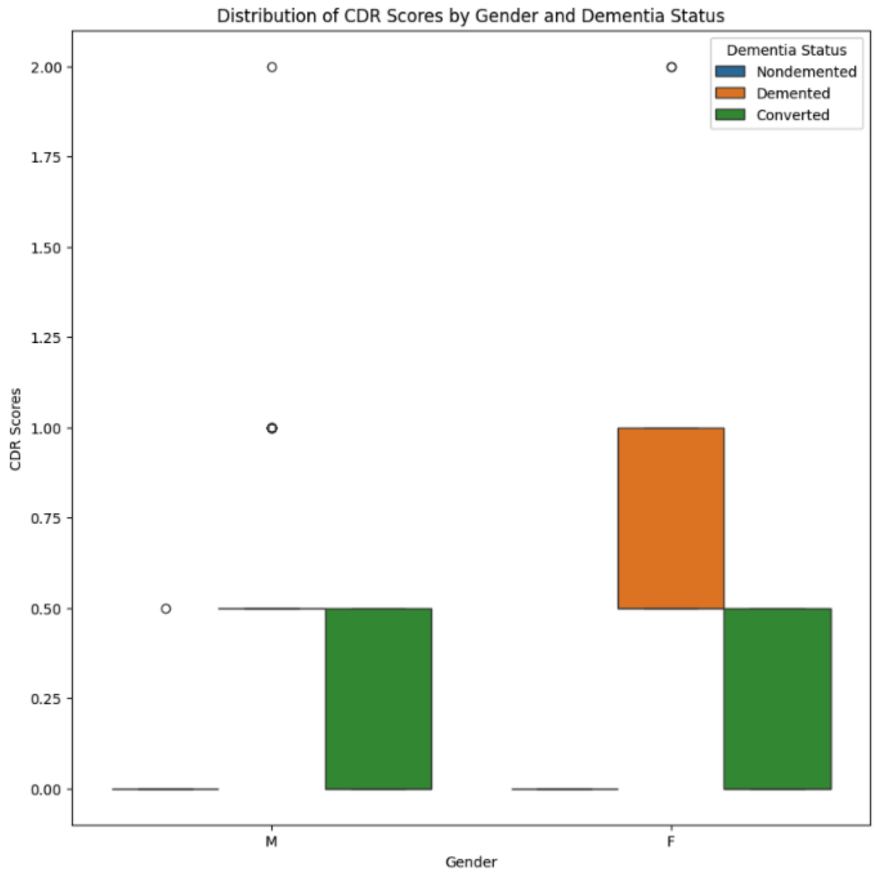

Figure 1. Distribution of CDR Scores by Gender and Dementia Status

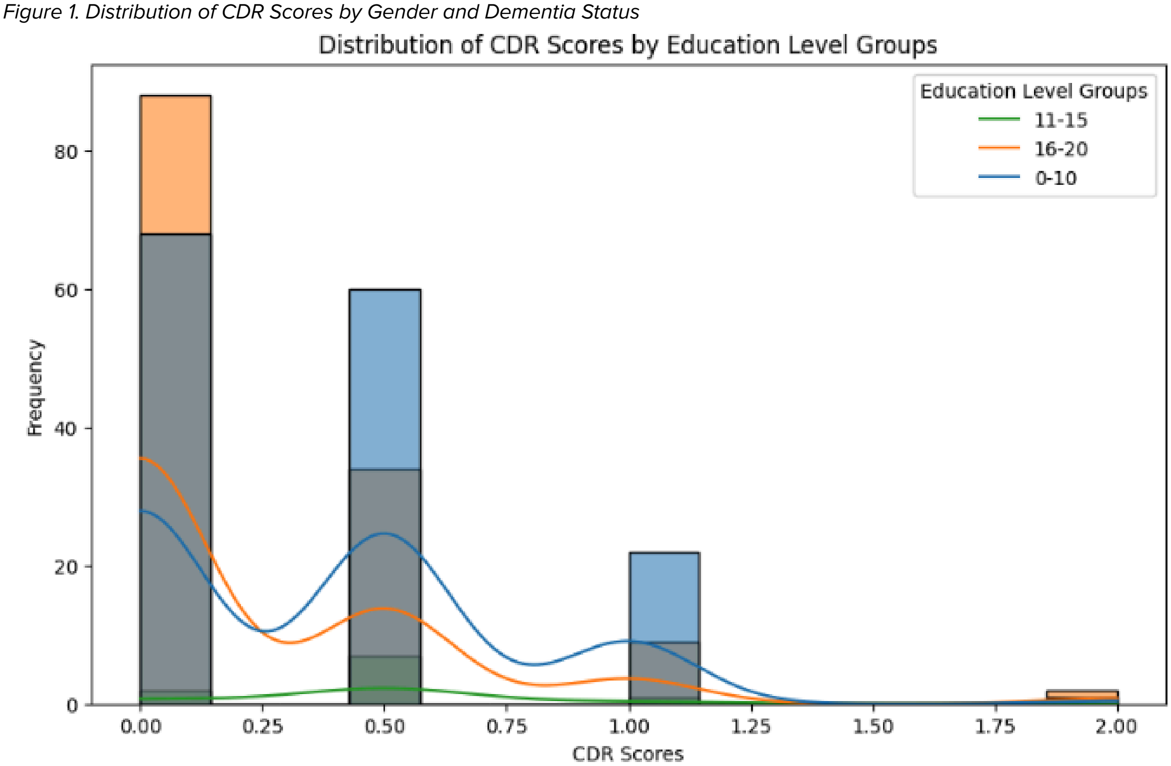

Figure 2. Distribution of CDR Scores by Education Level Groups

Quantitative Analysis: Mixed-Effects ANOVA Models

To address the research questions, I conducted individual ANOVA tests and interaction analyses using mixed-effects models:

Age Analysis: ANOVA for age revealed a significant effect on CDR scores ($F(37, 256) = 1.625$, $p = 0.0166$, $\eta^2 = 0.190$).

	ddof1	ddof2	F	p-unc	np2
age	37	256	1.625	0.0166	0.190

Gender Analysis: ANOVA for gender demonstrated associations between gender and CDR scores ($F(1, 292) = 9.102$, $p = 0.003$, $\eta^2 = 0.030$).

	ddof1	ddof2	F	p-unc	np2
gender	1	292	9.102	0.003	0.030

Dementia Status Analysis: ANOVA for dementia status showed significant differences in CDR scores among different dementia statuses ($F(2, 291) = 326.532$, $p = 4.302e-69$, $\eta^2 = 0.692$).

	ddof1	ddof2	F	p-unc	np2
dementia_status	2	291	326.532	4.302	0.692

Interaction analyses were performed to explore potential interactions between age, gender, and dementia status on CDR scores. The interaction between age and gender showed notable effects on CDR scores, while the interaction between age and dementia status provided insights into age-related differences in CDR scores based on dementia status.

Correlation Analysis

A correlation matrix was computed to examine relationships between CDR scores, age, gender, and dementia status. The matrix revealed correlations between these variables, informing subsequent analyses.

	CDR	age
CDR	1.000	0.011
age	0.011	1.000

Statistical Power Analysis for t-tests

To determine the required sample size for detecting a specified effect size in a two-sample t-test scenario, I conducted a statistical power analysis. Assuming a mean difference (effect size) of 0.7 in CDR scores between groups, with a desired power of 0.91 and a significance level of 0.05, the analysis estimated a required **sample size of approximately 46 participants per group** to achieve adequate power.

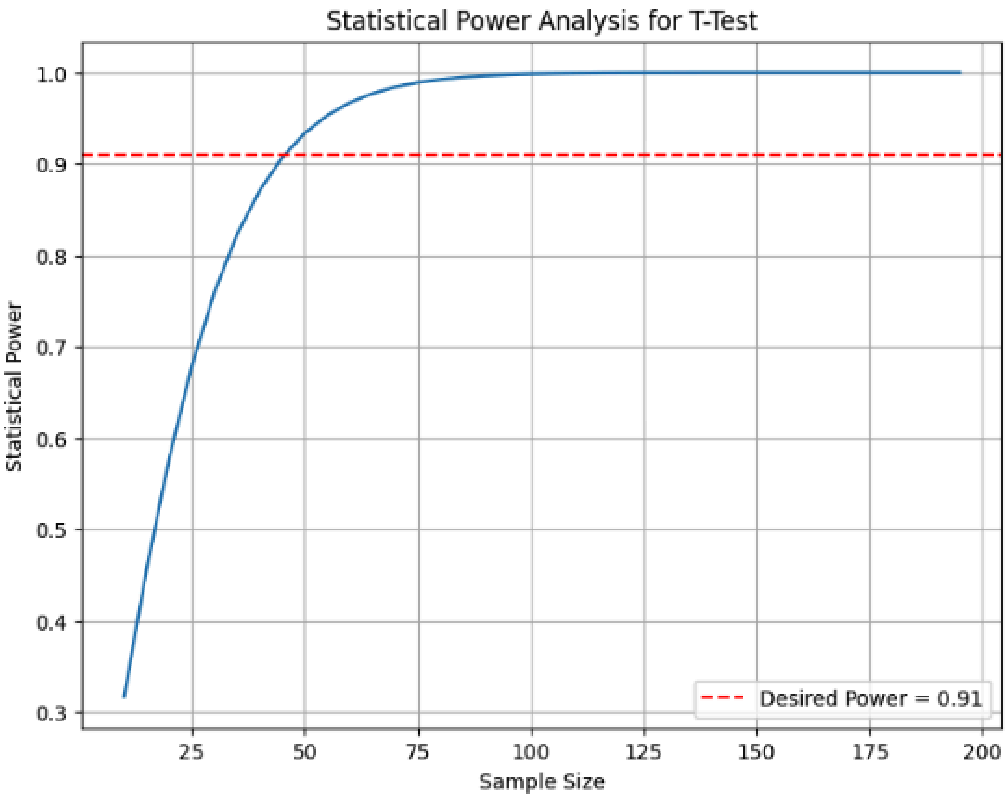

Figure 3. Statistical Power Analysis for T-Test

Analysis Discussions and Conclusion:

The descriptive statistics and boxplots demonstrated that CDR scores were significantly higher among patients with dementia compared to those without dementia, indicating that CDR scores effectively differentiate between patients with and without dementia based on demographic factors. Although not explicitly tested in this analysis, examining the interaction effect between education level and dementia status on CDR scores could reveal whether the impact of dementia on CDR scores varies across different levels of education, requiring further specific testing of this interaction. The individual ANOVA analyses indicated that age significantly influences CDR scores. Moreover, exploring interaction effects between age and gender, as well as age and dementia status, can provide insights into how age modifies the relationship between CDR scores and these demographic factors. Positive correlations observed between age and CDR scores suggest that age plays a role in influencing dementia-related outcomes.